



Effective methodologies in statistical inference with low sample size data

著者	矢田 和善
内容記述	Thesis (Ph. D. in Science)--University of Tsukuba, (A), no. 5255, 2010.3.25 Includes bibliographical references
発行年	2010
URL	http://hdl.handle.net/2241/105479

氏 名 (本籍)	矢 ^や 田 ^た 和 ^{かず} 善 ^{よし} (島 根 県)		
学 位 の 種 類	博 士 (理 学)		
学 位 記 番 号	博 甲 第 5255 号		
学位授与年月日	平成 22 年 3 月 25 日		
学位授与の要件	学位規則第 4 条第 1 項該当		
審 査 研 究 科	数理物質科学研究科		
学 位 論 文 題 目	Effective Methodologies in Statistical Inference with Low Sample Size Data (小標本データによる統計的推測の有効的方法論)		
主 査	筑波大学教授	博士 (理学)	青 嶋 誠
副 査	筑波大学教授	理学博士	坪 井 明 人
副 査	筑波大学教授	理学博士	笠 原 勇 二
副 査	筑波大学准教授	博士 (理学)	小 池 健 一

論 文 の 内 容 の 要 旨

本学位論文は、3つの章から構成される。

第1章は、統計的標本抽出の理論において、未解決とされていた2段階標本抽出の漸近一致性について、2次の漸近一致性をもつ解を与えることに成功している。各種推測の問題への応用を考え、特に、ジェネリック医薬品の承認要件における生物学的同等性の検定に、有意水準と検出力を同時に満たし、臨床試験の実施期間短縮とコスト削減を目的とした、新しい方法論を与えている。

第2章は、多次元統計的推測に2重縮小という新たな概念を導入し、高次元データの領域推定に、標本数を大幅に節約してもなお精度保証を可能にする方法論を与えている。2重縮小という概念に基づく対象領域の特定化は、定理で述べられている主張以上に、データ解析の数値が、実質的な有効性を示している。回帰分析などの多次元統計解析の各種分析理論への応用も与えている。

第3章は、本学位論文の核となる章である。高次元小標本データにおける統計的推測に、本学位論文で新たに開発された理論と方法論が凝縮されている。従来の多次元統計解析の理論には、データの次元数と標本数の大小関係に不文律がある。高次元小標本においては、多次元統計解析の理論は、次元の呪いによって破綻する。大標本漸近理論は使用できず、母集団分布に正規分布を仮定することは、もはや合理的といえない。高次元小標本のデータには、何らかのノンパラメトリックな方法論と、大標本漸近理論に替わる新たな漸近理論の構築が必要である。本学位論文では、高次元小標本データ特有のある幾何学的構造に注目し、高次元(小標本)漸近理論を整備して、クロスデータ行列法という新しいノンパラメトリック法を提唱し、潜在空間の次元数の推定に始まり、固有値・固有ベクトル・主成分スコアなどの各種特徴量の統計的推測について、新しい理論を構築している。

審 査 の 結 果 の 要 旨

マイクロアレイデータなど、近代科学で扱うデータの一つの特徴は、次元数が標本数よりも遥かに大きな

高次元小標本 (HDLSS) にある。HDLSS における理論は完全に整備されているとはいえず、理論の開拓と新たな方法論の開発が急務とされている。

HDLSS において理論を構築することの困難さは、高次元データ空間を埋め尽くす巨大なノイズとモデルの特異性、それらが織り成す不可思議な「次元の呪い」に起因する。

矢田氏は、従来の理論や方法論には見られない独創的なアイデアと解析法を考案し、この未開拓であった研究領域に、主に、以下の3つの重要な仕事をした。

- (1) 従来の統計学の理論と方法論が HDLSS に導く結果は、推測の一致性すら有しないことの証明。
- (2) HDLSS 特有の幾何学的構造の発見と、次元の呪いを理論的に捉えることの成功。
- (3) HDLSS のノイズの挙動を解析するための、高次元 (小標本) 漸近理論の開発と、ノイズの大きさの精密な見積り、そして、新しい統計的推測「クロスデータ行列法」の提唱とその一致性の証明。

高次元データは本来、ノイズに埋もれているものである。高次元空間からの産物は、無視できないほど大きなノイズを含んだものになっている。ノイズの影響を無視して、大標本漸近理論をもとにした従来型の多次元統計解析を HDLSS に使うと、誤った結果が導かれる。それを理論的に証明するには、新しく高次元 (小標本) 漸近理論を整備する必要がある。それが本学位論文で与えられている。例えば、高次元空間の潜在空間は、小標本では定まらない。従来型の多次元統計解析では、固有ベクトルの推定が 90 度方向性を誤り、一意に方向を定められない。本学位論文による理論と方法論を使えば、その方向も定めることができる。

矢田氏の功績は、上述したように、HDLSS に特有な幾何学的構造を発見し、ノイズが占める大きさと挙動を精密に解析するための高次元 (小標本) 漸近理論という新しい理論を展開して、それに基づいた統計的推測の有効な方法論を考案したことにある。提唱された方法論は、従来の統計学にはない独創的なアプローチから生み出されており、理論に裏打ちされているだけでなく、実際のデータ解析において有効であることが示されている。

数理統計学の未開拓領域に、インパクトを与える重要な結果が、理論と方法論の両方にバランスよく、かつ重厚に開発されている。未開拓領域を、これだけ整備した研究業績は、矢田氏の独創的なアイデアとそれを証明する解析学の高い能力があってこそ、成しえた結果であろうと思われる。

論文審査ならびに最終試験の結果に基づき、著者は博士 (理学) の学位を受けるに十分な資格を有するものと認める。